2.1 The General Linear Generator

Remember that a general linear generator is characterized by

- a ring $R$ and an $R$-module $M$ as external parameters,
- a linear map $A: M \to M$ as internal parameter,
- a sequence of vectors $x_n \in M$ as states and output elements,
- a vector $x_0 \in M$ as initial state,
- a recursive formula $x_n = Ax_{n-1}$ for $n \geq 1$ as state transition.

Remark (the trivial case): If $A$ is known, then from each member $x_r$ of the output sequence we may predict all of the following members $(x_n)_{n>r}$. Therefore this case lacks cryptological relevance. Note that calculating the sequence backwards, that is $x_n$ for $0 \leq n < r$, is uniquely possible only if $A$ is injective. But this effect doesn’t rescue the cryptologic value of the generator. For simplicity in the following we usually treat forwards prediction only, assuming that an initial chunk $x_0, \ldots, x_{k-1}$ of the output sequence is known. However we should bear in mind that also backwards “prediction” might be an issue.

Assumption for the following considerations: $R$ and $M$ are known, $A$ is unknown, and an initial segment $x_0, \ldots, x_{k-1}$ is given. To avoid trivialities we assume $x_0 \neq 0$. The prediction problem for this scenario is: Can the attacker determine $x_k, x_{k+1}, \ldots$?

Yes she can, provided she somehow finds a linear combination

$$x_k = c_1 x_{k-1} + \cdots + c_k x_0$$

with known coefficients $c_1, \ldots, c_k$. For then

$$x_{k+1} = A x_k = c_1 A x_{k-1} + \cdots + c_k A x_0$$
$$= c_1 x_k + \cdots + c_k x_1$$
$$\vdots$$
$$x_n = c_1 x_{n-1} + \cdots + c_k x_{n-k} \quad \text{for all } n \geq k,$n

and by this formula she gets the complete remaining sequence—without determining $A$ (!). But how to find such a linear combination?

A simple example is periodicity: $x_n = x_{n-k}$ for all $n \geq k$. Linear algebra provides a more general solution. In the present abstract framework it is natural to assume $M$ as Noetherian (usually the “proper” generalization of a finite-dimensional vector space). Then the ascending chain of submodules

$$Rx_0 \subseteq Rx_0 + Rx_1 \subseteq \cdots \subseteq M$$
is stationary: there is an \( r \) with \( x_r \in Rx_0 + \cdots + Rx_{r-1} \). And this yields the linear relation we need; of course it is useful only when we succeed with explicitly determining the involved coefficients. Note that a finite module \( M \)—that we usually consider for random generation—is trivially Noetherian.

By this consideration we have shown:

**Proposition 4** (Noetherian principle for linear generators) Let \( R \) be a ring, \( M \), an \( R \)-module, \( A : M \to M \) linear, and \( (x_n)_{n \in \mathbb{N}} \) a sequence in \( M \) with \( x_n = Ax_{n-1} \) for \( n \geq 1 \). Then for \( r \geq 1 \) the following statements are equivalent:

(i) \( x_r \in Rx_0 + \cdots + Rx_{r-1} \).

(ii) There exist \( c_1, \ldots, c_k \in R \) such that \( x_n = c_1 x_{n-1} + \cdots + c_r x_{n-k} \) for all \( r \geq k \).

If \( M \) is Noetherian, then an \( r \) with (i) and (ii) exists.

But how to explicitly determine the index \( k \) and the coefficients \( c_1, \ldots, c_k \)? Of course this can work only for rings \( R \) and modules \( M \) that admit explicit arithmetic operations.

In the following our main examples are: \( R = K \) a finite field, or \( R = \mathbb{Z}/m\mathbb{Z} \) a residue class ring of integers. In both cases we have a-priori knowledge on the number of true increments in the chain of submodules; that is, an explicit bound for \( r \). If for example \( R \) is a field, then the number of proper steps is bounded by the vector space dimension \( \dim M \). In the general case we have:

**Proposition 5** (Krawczyk) Let \( M \) be an \( R \)-module, and \( 0 \subset M_1 \subset \cdots \subset M_l \subset M \) be a properly increasing chain of submodules. Then \( 2^l \leq \#M \).

This result is useful only for a finite module \( M \). However this is the case we are mainly interested in when treating congruential generators. Then we may express it also as \( l \leq \log_2(\#M) \). This is not too bad compared with the case field/vector space, both finite: \( l \leq \dim(M) \leq \log_2(\#M)/\log_2(\#R) \).

Proof. Let \( b_i \in M_i - M_{i-1} \) for \( i = 1, \ldots, l \) (where \( M_0 = 0 \)). Then the subset

\[ U = \{ c_1 b_1 + \cdots + c_l b_l \mid \text{all } c_i = 0 \text{ or } 1 \} \subseteq M \]

consists of \( 2^l \) distinct elements. For if two of these expressions would represent the same element, their difference would have the form

\[ e_1 b_1 + \cdots + e_t b_t = 0 \quad \text{with } e_i \in \{0, \pm 1\}, \quad e_t \neq 0, \]

for some \( t \) with \( 1 \leq t \leq l \). From \( e_t = \pm 1 \in R^\times \) we would derive the contradiction \( b_t = -e_t^{-1}(e_1 b_1 + \cdots + e_{t-1} b_{t-1}) \in M_{t-1} \). Hence \( \#M \geq \#U = 2^l \). \( \square \)